John von Neumann Institute for Computing

Milestoning

R. Elber, A. West

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 13-22, 2008.

http://www.fz-juelich.de/nic-series/volume40

Milestoning

Ron Elber and Anthony West

Department of Chemistry and Biochemistry, ICES, C0200
The University of Texas at Austin, Austin TX 78712, USA
E-mail: ron@ices.utexas.edu

Milestoning is a theory and an algorithm to compute kinetics and thermodynamics of complex molecular systems. It makes it possible to study general processes on rugged energy landscapes on timescales not approachable by straightforward Molecular Dynamics (microseconds and milliseconds). The algorithm is based on monitoring progress along a set of discrete states (Milestones) using short-time microscopic trajectories that capture local dynamics. These discrete states can be (for example) hypersurfaces perpendicular to a reaction coordinate. Transition times between Milestones are recorded to produce Local-First-Passage-Time-Distributions (LFPTD). The theory is based on a non-Markovian integral equation for the probability flow between Milestones. The integral equation is equivalent to the Generalized Master equation. No specific model is assumed for the microscopic dynamics. The theory uses the LFPTD to compute the overall kinetic and thermodynamic. Complex transitions in proteins were investigated (allosteric transition in Scapharca hemoglobin, the recovery stroke in myosin). In the present review only the simple example of alanine dipeptide is discussed.

1 Introduction

Atomically detailed simulations provide useful information on biomolecular processes using a single unified model. Specifically, Molecular Dynamics (MD) algorithms are available to compute efficiently thermodynamic and equilibrium behavior. However, MD is limited when studying non-equilibrium processes and kinetics. Straightforward and typical trajectories of condensed phase systems rarely exceed hundreds of nanoseconds, far too short to investigate the kinetics of many interesting biophysical systems. Examples are of conformational transitions, ion permeation, protein folding and more. Extending the time scale of molecular simulations is therefore an important research direction and has attracted the attention of many investigators.

It is useful to classify processes of long time dynamics into two categories: Dynamics which are (i) activated or (ii) diffusive (figure 1). Significant progress has been made in algorithm design and theory development for activated processes[2, 14, 8, 6, 11]. In activated processes rare short time trajectories pass over significant free energy barriers and determine the overall kinetics. Progress has been slower for diffusive processes (or a mixture of activated and diffusive processes) in which the times of the individual transitional trajectories are intrinsically long. Diffusion on rugged energy landscapes is not necessarily associated with a narrow transition domain between stable states. A narrow transition domain is typical in activated processes and facilitates the use of short time trajectories to probe reactive events. If we probe an activated system at different time slices, in the majority of the observations we do not observe something new. The system remains in the reactant state until a rapid (but rare) transition is initiated to the product state. In contrast, probing diffusive processes show spatial progress in sequential observations. Milestoning is a theoretical and computational approach that aims at diffusive or mixed processes. Nevertheless, it can

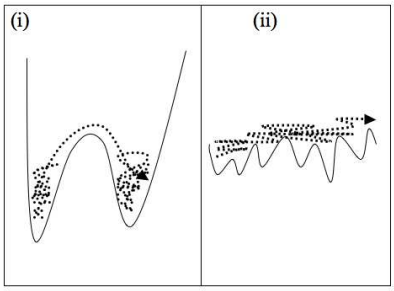

Figure 1. A schematic representation of an (i) activated and (ii) diffusive energy landscape that leads to different (corresponding) types of dynamics.

also handle activated processes and therefore suggests a uniform technology for the two types of dynamics.

A conceptual approach to long time dynamics is that of coarse graining in space and time. Indeed a number of groups have followed this idea, and have proposed fitting parameters of a kinetic model[13,1] or of the diffusion equation[16] based on atomically detailed simulations. For example, it is assumed that rate constants (exponential relaxations in time) describe transitions between the states of a Master equation. Power law and stretched exponential kinetics were found in biophysical kinetics[7]. Moreover, there is no rigorous mapping from an atomically detailed description of the system to a diffusion equation and the decision of what exactly to fit is not unique.

In contrast to the phenomenological modeling of the Master equation there is a rigorous approach to spatial and temporal coarse graining by Zwanzig and Mori. It is the Generalized Langevin Equation[17] or equivalently the Generalized Master Equation[10]. A memory kernel (and not a rate constant) describes the impact of the "bath". Unfortunately, the numerical calculations of the memory kernel of the Generalized Master Equation are difficult, motivating the use of the simpler and less rigorous Master Equation. At the core of the Milestoning approach one finds an algorithm to circumvent the difficulty in computing the memory kernel. The function we compute is formally equivalent but easier to estimate numerically than the rate kernel. Therefore the Milestoning approach is equivalent to the Generalized Master Equation and is based on a rigorous theory of non-equilibrium processes.

2 Milestoning

2.1 Theory

We describe below the approach of Milestoning. The Milestoning theory follows from an intuitive expression in which we capture the characteristics of the microscopic dynamics in a non-Markovian kernel, and then solve the dynamics within the assumption of local equilibrium.

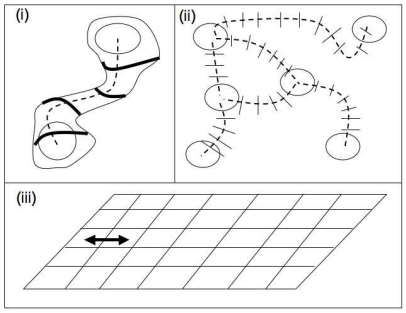

Figure 2. Examples of concrete realizations of the Milestoning idea: (i) Hypersurfaces (thick lines) perpendicular to a reaction coordinate (dashed line), (ii) hypersurfaces along reaction coordinates (thick dashed line) connecting local free energy minima between local free energy minima, and (iii) a grid over a few collective variables (illustrated for two). The Milestones are the boundaries of the squares. A potential transition at one interface it denoted with the double-ended arrow.

Let the vector describing the atomically detailed system be $X \in R^N$. We have at our disposal an operator D that generates trajectories such that $D\left[X\left(0\right),t\right] = X\left(t\right)$ where t is the time. Depending on the type of dynamics at hand X is a coordinate or a phase-space vector. The numerical application of D is usually made in small time steps and is repeated many times to produce a long time trajectory $X\left(t\right) = D^N\left[X\left(0\right),\Delta t\right], \Delta t = t/N$. The generation of the trajectories is by far the most expensive calculation at hand (also in Milestoning). Typically we will be interested in an ensemble of reactive trajectories that transition from a state of reactant to a state of product. The direct sampling of the transition (i.e. computing $X\left(t\right) \in$ product while $X\left(0\right) \in$ reactant) is assumed too expensive to pursue with direct use of D (of course, if it is not too expensive an exact solution of the problem is always better than a solution based on approximations or physical assumptions).

We now discuss the coarse grained model that we use to analyze the microscopic data and extend the time scale of the simulation. We partition the space and define discrete states s. Two of these states $s = 0$ and $s = L$ are the reactants and products respectively. Other $L-1$ states $s = 1, ..., L-1$ are called "Milestones". A realization of the space partition and "Milestones" is of (i) hypersurfaces perpendicular to a reaction coordinate, (ii) local free energy minima, and (iii) a grid over a reduced set of collective variables (figure 2). The examples discussed in the present manuscript (and in our published papers[6,4,15]) are of hypersurfaces perpendicular to reaction coordinates. However, the formulation below applies just as well to (ii) and (iii).

The probability of being at Milestone s at time t is $P_s\left(t\right)$. To compute $P_s\left(t\right)$ we derive from atomically detailed simulations the transition probability density $K_{s,s'}\left(\tau\right)$. It is the probability of making a transition from state s into state s' after an "incubation/waiting" time τ in the state s. The incubation time captures memory effects in which the transition probability between s and s' depends on the time spent already in s. It is the only microscopically derived function that we need for Milestoning. Note that we assume that the transition probability is independent of the absolute time. This assumption is not valid in systems that strongly deviate from equilibrium or a stationary state. We argued and illustrated[4,15] that it is much easier to compute the above matrix than to perform the complete simulations from reactants to products. Scaling arguments in support of the expected speed up are presented in the section Algorithm. This matrix is finally used in a probabilistic non-Markovian framework to obtain the overall kinetics of the system.

For further development it is convenient to define another function $Q_s\left(t\right)$. It is the

probability density that a trajectory will make a transition into s at time t. With $K_{s,s'}(\tau)$ and the definition of $Q_s(t)$ at hand, equation (1) below simply balances transition probabilities. The system is initiated at zero time and starting probabilities are injected into the Milestones. At later times we consider transitions between the states. To make a transition into s from one of the nearby s' it is necessary to transition first to s', wait (or incubate) at s' for time τ and then transition into s. The probability density of making a transition into s' at time $t-\tau$ is $Q_{s'}(t-\tau)$. The probability density of making a transition from s' to s after waiting time τ is $K_{s,s'}(\tau)$. Finally, a summation over all states s' that are directly connected to s, and over all incubation times τ gives equation (1) below.

$$Q_s(t) = \delta\left(t-0^+\right) P_s(0) + \int_0^t \sum_{s'} K_{s,s'}(\tau)\, Q_{s'}(t-\tau)\cdot d\tau \tag{1}$$

It is a matrix-vector equation in s space and an integral equation in time. The unknown is the vector of functions $Q_s(t)$. This equation is called in physics CTRW (Continuous Time Random Walk) and was used in phenomenological modeling of transport[9]. We are connecting this equation with atomically detailed simulations and use it to study long time phenomena. Microscopic dynamics is used to compute $K_{s,s'}(\tau)$.

Our interest focuses on $P_s(t)$, the probability of being at s at time t. With $Q_s(t)$ determined from equation (1) we write $P_s(t)$ as the integral of probabilities to make a transition into s at an earlier time t' and to remain at s (avoid transitions to other states s') until time t. Summation over all channels s' gives

$$P_s(t) = \int_0^t Q_s(t') \left[1 - \sum_{s'} \int_0^{t-t'} K_{s',s}(\tau)\cdot d\tau\right] dt' \tag{2}$$

Equations (1) and (2) are very general. They do not assume concr ete dynamics or mechanism, only that a transition matrix (with memory) can be defined (and computed) between states that represent the system. The solution of (1) and (2) is now a topic in applied mathematics and was obtained with different approaches[6,12,15]. The most obvious one is to solve the integral equation (equation (1)) by small time steps. Since the number of degrees of freedom was greatly reduced and the functions considered are much smoother with respect to time compared to MD, the computational efforts are still negligible compared to the calculations of the trajectories.

$$\begin{aligned} Q_s(0) &= P_s(0)/\Delta t \\ Q_s(\Delta t) &= Q_s(0) + \sum_{s'} K_{s,s'}(\Delta t)\cdot Q_{s'}(0)\cdot \Delta t \\ Q_s(2\Delta t) &= Q_s(0) + \sum_{s'} \left[K_{s,s'}(\Delta t)\cdot Q_{s'}(\Delta t) + K_{s,s'}(2\cdot\Delta t)\cdot Q_{s'}(0)\right] \\ &\ldots\ldots \end{aligned} \tag{3}$$

Another solution is based on Laplace transforms on (1) and (2) and algebraic manipulation of the transforms[12]. We quote only one result. Define the off diagonal matrix $\left(\hat{K}\right)_{s,s'}(t) = K_{s,s'}(t), s \neq s'$, the time integral $\langle f\rangle \equiv \int_0^\infty f(t)\, dt$ and an average over an ensemble of trajectories by $\bar{f}$. A useful measure for the kinetic properties of the system is the overall first passage time. It is defined as the time required for a trajectory initiated

at the reactant to reach the product state for the first time. The ensemble average of the overall first passage time $\bar{\tau}$ is given by

$$\bar{\tau} = I \cdot \left\langle \tau \cdot \hat{K}_{s,s'}(\tau) \right\rangle \left[I - \left\langle \hat{K}_{s,s'}(\tau) \right\rangle \right]^{-1} \cdot \varepsilon_i \tag{4}$$

where I is the identity matrix and ε_i is a unit vector in the direction of the initial Milestone i. The last Milestone is set to be absorbing.

2.2 Algorithm

It is important to emphasize that the theory described in the previous section is "equation free". We do not assume Langevin, Brownian or Newtonian mechanics. All reasonable forms of dynamics can be used to generate numerical values of the transition matrix. We compute the transition matrix as discussed below and we do not fit parameters to a particular coarse-grained dynamical model.

The only task of the microscopic dynamics is to compute the transition matrix $K_{s,s'}(\tau)$. This matrix is used in equation (1-4) to determine the overall rate and the evolution of the system in time. To facilitate the calculation and make it highly efficient compared to straightforward MD we restrict our attention to systems that satisfy the following requirements:
Condition (i): The system is in a stationary state (called ρ_s for Milestone s). Only a few variables may be left non-stationary.

Condition (ii): Trajectories that arrive at Milestone s' are distributed in the hyperplane according to the stationary distribution $\rho_{s'}$.

If the system is in equilibrium, or sufficiently close to it then assumption (i) is obviously satisfied. If the equilibrium is canonical then the probability density is $\rho_s \simeq \exp[-\beta U(X)]$, $X \in s$. This is the weight that was used in the alanine dipeptide example discussed in this paper. Alternatively, it is possible to have the system at a stationary (time-independent) but non-equilibrium state. An example is of a stationary flow of liquid in a confined environment. The second condition is subtler than the first. It requires the states s to be well separated in time so that the distribution of trajectories arriving to s' is the same as the stationary distribution of condition (i), i.e. $\rho_{s'}$. For example, this condition is satisfied if the time scale for the transition is longer that the time scale to reach local equilibrium at each of the Milestones. In practice it is possible to monitor and vary the transition time scales by placing the Milestones sufficiently apart with the second condition in mind. Of course, the choice of Milestoning should make the transition times between Milestones much shorter than the overall time scale of the reaction. Otherwise, Milestoning is not advantageous to MD.

The calculation of $K_{s,s'}(\tau)$ is done by sampling trajectories between a specific pair of states s and s'. The trajectories are initiated at s according to the stationary distribution ρ_s and their termination times at s' are recorded. These distributions are binned to estimate the probability that the system will transition between s and s' after incubation time τ. They are called the Local-First-Passage-Time-Distribution (LFPTD), and are also the matrix elements $K_{s,s'}(\tau)$. The LFPTD is normalized such that $f_{s,s'} = \int_0^\infty K_{s,s'}(\tau)\, d\tau$ is the

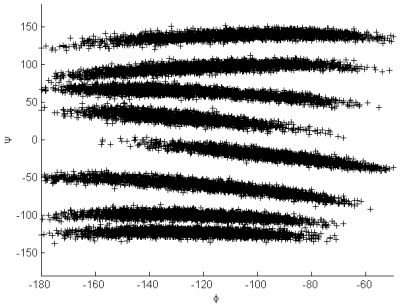

Figure 3. Sampling equilibrium distributions in the Milestones (hyperplanes perpendicular to the reaction coordinate) of the transition in alanine dipeptide. The distributions are projected onto a two-dimensional surface of internal coordinates (ϕ, ψ). Note that all simulation includes all coordinates of the dipeptide's atom and the coordinates of periodic box, solvating water molecules.

fraction of trajectories that were initiated at s and terminate at s' $\left(\sum_{s'} f_{s,s'} = 1\right)$.

What is the gain in studying the kinetic with Milestones instead of using straightforward MD? This important question was discussed extensively[15]. Practical gains were clearly illustrated in the examples of 4, 15. For completeness we quote an argument for computational speedup expected for diffusive processes[15]. The required computational resources are proportional to the number of force evaluations and therefore to the lengths of trajectories that reach the product state. Consider a reaction in which the system diffuses freely for a length L. The time to react using straightforward trajectories is proportional to L^2. In Milestoning we chop the complete length L to (say) N pieces. The time to diffuse through one piece is $(L/N)^2$. There are N pieces and therefore the time to destination in Milestoning is $(L/N)^2 \cdot N = L^2/N$. We obtain a speed up proportional to the number of Milestones used. Exponential speedup for systems with free energy barrier can be illustrated as well[15].

In the next paragraph we discuss a concrete example. We consider progress along a reaction coordinate measured by passing hypersurfaces (Milestones) orthonormal to it. In figure 2.ii we sketch a terminating trajectory between Milestones s and $s-1$. An ensemble of such trajectories is used to compute the probability densities $K_{s,s-1}(\tau)$ and the overall rate according to equation (4).

2.3 Example: Folding of a Solvated Dipeptide

We review a detailed calculation of the kinetics of a solvated dipeptide[15]. All the calculation described below were performed with the MOIL program which is in the public domain[5]. The Milestones were constructed from adiabatic energy surface of the ψ dihedral angle. The energy of the peptide in vacuum was minimized with the ϕ dihedral angle constrained to 150^o the ψ dihedral angle constrained to values between -180^o and +180^o degrees with a step size of 2.5 degrees. The minimizations provided a total of 144 structures (and potential Milestones) that we denote by $s = 1, ..., 144$. Milestones, which are hyperplanes perpendicular to the reaction coordinate, are defined by the coordinate of the minimized structures X_s, and the numerically estimated normals to the hyperplanes $q_s \equiv \frac{X_{s+1} - X_{s-1}}{|X_{s+1} - X_{s-1}|}$.

The different configurations were solvated in water boxes of volume (20Å)^3 and 248 water molecules. Molecular dynamics simulations constrained to each of the hyperplanes[3]

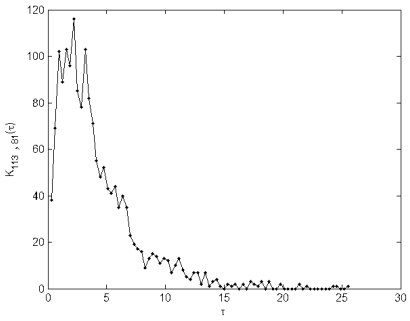

Figure 4. A typical example for a local first passage time distribution in a solvated alanine dipeptide. The example below is for a seven milestone partitioning of the reaction coordinate. The computed transition is from Milestone six to five (the indices below refer to the total of 144 Milestones that we considered using numerous variations[15]. The calculation with seven Milestones gives the correct rate. Even on this very simple example we obtain significant speedup. The average transition time between any two Milestones in the seven Milestone run is about 3.6 picoseconds. A cost of a single trajectory that passes all Milestones from the first to the last will be $3.6 \times 6 = 21.6$ picoseconds. This is about 3 times shorter than the correct first passage time of about 60 picoseconds.

were used to sample equilibrium distributions in the Milestones. In figure 3 we display a projection of the simulations on a (ϕ, ψ) map. Since the hyperplanes are defined in Cartesian space the projection on the space of internal coordinates show significant "width". We verified that the simulations create exact hyperplanes in Cartesian space.

M	$\bar{\tau}(ps)$	$\tilde{\tau}(fs)$
144	500(3.1)	31.2
74	261(1.2)	57.7
73	330(1.6)	58.3
37	104(0.63)	129
19	62(0.47)	373
11	53(0.50)	1,305
7	62(0.73)	3,581
5	68(0.93)	10,902
3	$64(1.04)$	-

Table 1. A summary of runs for alanine dipeptide conformational transition from an alpha helix to an extended chain conformation. The first column is the number of Milestones. The case of three Milestones (last row) is exact. The second column is the estimated overall first passage time and the third column the average LFPTD ($\tilde{\tau}$ - equation (5)). The runs differ in the number of Milestones used. When the number of Milestones is larger then the speedup is more significant. However if the number of Milestones is too large the local equilibrium assumption is violated and the rate is inaccurate. We compare the velocity decorrelation time (400 femtoseconds) with the average transition time between Milestones. If the number of Milestones is larger than 19 then the LFPTD is shorter than 400 femtoseconds and the rate is wrong. If the numbers of Milestones is smaller than 19 the results are quite accurate.

In the next step we initiate MD trajectories starting from the configurations sampled at each of the Milestones, s. The trajectories were integrated to termination at Milestones $s \pm 1$. The distributions of termination times are the elements of the transition matrix $K_{s,s\pm1}(t)$. A typical result is shown in figure 4. Initially we have used 144 Milestones. However, in this case the termination times are very short and they do not satisfy the condition about relaxation to equilibrium (condition (ii)). A useful test for relaxation is the velocity decorrelation time that we found in that case to be about 400 femtoseconds (figure

8 of 15). A typical transition time between Milestones, estimated as the following average

$$\tilde{\tau} \equiv \frac{1}{L}\sum_{s}\bar{\tau}_s = \frac{1}{L}\int_0^{\infty}\tau_s \cdot \left(K_{s,s+1}\left(\tau\right) + K_{s,s-1}\left(\tau\right)\right)d\tau \tag{5}$$

The typical transition time must be larger than the velocity de-correlation time. In table 1 we show that this condition is satisfied only for a number of milestones smaller than 19. All the calculations with a number of milestones smaller or equal to 19 approximate well the exact rate.

This calculation illustrates that Milestoning provides accurate results for a non-trivial but exactly solvable model. It was also shown that it is significantly more efficient. Comparing directly simulation time, we have concluded in our publication that in this example Milestoning was more efficient by a factor of about 9. In other calculations of larger more complex systems (e.g. the allosteric transition in Scapharca hemoglobin), the speedup was a factor of about 1,000.

References

1. Chodera, J. D., W. C. Swope, et al. (2006). *Long-time protein folding dynamics from short-time molecular dynamics simulations.* Multiscale Modeling & Simulation **5**(4): 1214-1226.
2. Dellago, C., P. G. Bolhuis, et al. (2002). *Transition path sampling.* Advances in Chemical Physics **123**: 1-78.
3. Elber, R. (1990). *Calculation of the potential of mean force using molecular dynamics with linear constraints – An application to a conformational transition in a solvated dipeptide.* Journal of Chemical Physics **93**(6): 4312-4321.
4. Elber, R. (2007). *A milestoning study of the kinetics of an allosteric transition: Atomically detailed simulations of deoxy Scapharca hemoglobin.* Biophysical Journal **92**(9): L85-L87.
5. Elber, R., A. Roitberg, et al. (1995). *Moil a program for simulations of macromolecules.* Computer Physics Communications **91**(1-3): 159-189.
6. Faradjian, A. K. and R. Elber (2004). *Computing time scales from reaction coordinates by milestoning.* Journal of Chemical Physics **120**(23): 10880-10889.
7. Frauenfelder, H., B. H. McMahon, et al. (2001). *The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin.* Proceedings of the National Academy of Sciences of the United States of America **98**(5): 2370-2374.
8. Hummer, G. and I. G. Kevrekidis (2003). *Coarse molecular dynamics of a peptide fragment: Free energy, kinetics, and long-time dynamics computations.* Journal of Chemical Physics **118**(23): 10762-10773.
9. Kenkre, V. M. and R. S. Knox (1974). Generalized Master Equation Theory of Excitation Transfer.
10. Mori, H., H. Fujisaka, et al. (1974). A new expansion of the master equation.
11. Ren, W., E. Vanden-Eijnden, et al. (2005). *Transition pathways in complex systems: Application of the finite-temperature string method to the alanine dipeptide.* Journal of Chemical Physics **123**(13).

12. Shalloway, D. and A. K. Faradjian (2006). *Efficient computation of the first passage time distribution of the generalized master equation by steady-state relaxation.* Journal of Chemical Physics **124**(5).
13. Sriraman, S., L. G. Kevrekidis, et al. (2005). *Coarse master equation from Bayesian analysis of replica molecular dynamics simulations.* Journal of Physical Chemistry B **109**(14): 6479-6484.
14. Voter, A. F., F. Montalenti, et al. (2002). *Extending the time scale in atomistic simulation of materials.* Annual Review of Materials Research **32**: 321-346.
15. West, A. M. A., R. Elber, et al. (2007). *Extending molecular dynamics time scales with milestoning: Example of complex kinetics in a solvated peptide.* Journal of Chemical Physics **126**(14).
16. Yang, S. C., J. N. Onuchic, et al. (2007). *Folding time predictions from all-atom replica exchange simulations.* Journal of Molecular Biology **372**(3): 756-763.
17. Zwanzig. R. (2001). Nonequilibrium statistical mechanics. Oxford, Oxford University Press.